Blocked and Multishift Variants of the QZ Algorithm for Computing Deflating Subspaces of Regular Matrix Pencils

Bo Kågström

Dept. of Computing Science and HPC2N
Umeå University, Sweden

Daniel Kressner

Institut für Mathematik
TU Berlin

DFG research center Berlin
mathematics for key technologies

23.03.2004

Introduction

Want to solve generalized eigenvalue problem for matrix pencil

$$A - \lambda B, \quad A, B \in \mathbb{R}^{n \times n}.$$

This consists of:

- Finding generalized eigenvalues λ:

$$\det(A - \lambda B) = 0.$$

- Finding right and left deflating subspaces $\mathcal{X}$ and $\mathcal{Y}$:

$$A\mathcal{X} \subseteq \mathcal{Y}, \quad B\mathcal{X} \subseteq \mathcal{Y}.$$

Introduction

Want to solve generalized eigenvalue problem for matrix pencil

$$A - \lambda B, \quad A, B \in \mathbb{R}^{n \times n}.$$

This consists of:

- Finding generalized eigenvalues λ:

  $$\det(A - \lambda B) = 0.$$

  (for nonsingular B : $\Leftrightarrow$ λ is an eigenvalue of $B^{-1}A$)

- Finding right and left deflating subspaces $\mathcal{X}$ and $\mathcal{Y}$:

  $$A\mathcal{X} \subseteq \mathcal{Y}, \quad B\mathcal{X} \subseteq \mathcal{Y}.$$

  (for nonsingular B : $\Leftrightarrow$ $\mathcal{X}$ is an invariant subspace of $B^{-1}A$)

Introduction

Want to solve generalized eigenvalue problem for matrix pencil

$$A - \lambda B, \quad A, B \in \mathbb{R}^{n \times n}.$$

This consists of:

- Finding generalized eigenvalues λ:

  $$\det(A - \lambda B) = 0.$$

  (for nonsingular B : $\Leftrightarrow$ λ is an eigenvalue of $B^{-1}A$)
- Finding right and left deflating subspaces $\mathcal{X}$ and $\mathcal{Y}$:

  $$A\mathcal{X} \subseteq \mathcal{Y}, \quad B\mathcal{X} \subseteq \mathcal{Y}.$$

  (for nonsingular B : $\Leftrightarrow$ $\mathcal{X}$ is an invariant subspace of $B^{-1}A$)

Assumption: $A - \lambda B$ is a regular pencil, i.e., $\det(A - \lambda B) \not\equiv 0$.

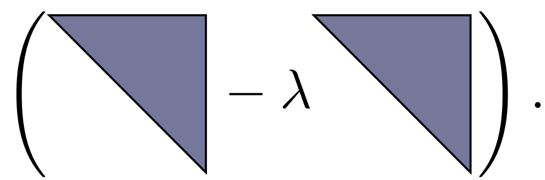
The Basic QZ Algorithm

Moler/Stewart '73: QZ generates a sequence of orthogonally equivalent matrix pencils:

$$(A_0 - \lambda B_0) := (A - \lambda B),\ (A_1 - \lambda B_1),\ (A_2 - \lambda B_2), \ldots .$$

Under suitable conditions (Watkins/Elsner '94) :

$$(A_i - \lambda B_i) \longrightarrow$$



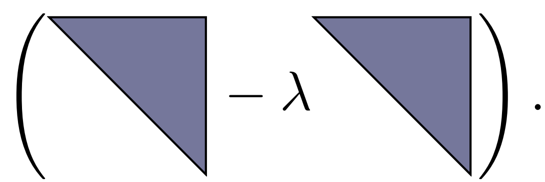
The Basic QZ Algorithm

Moler/Stewart '73: QZ generates a sequence of orthogonally equivalent matrix pencils:

$$(A_0 - \lambda B_0) := (A - \lambda B),\ (A_1 - \lambda B_1),\ (A_2 - \lambda B_2), \ldots .$$

Under suitable conditions (Watkins/Elsner '94) :

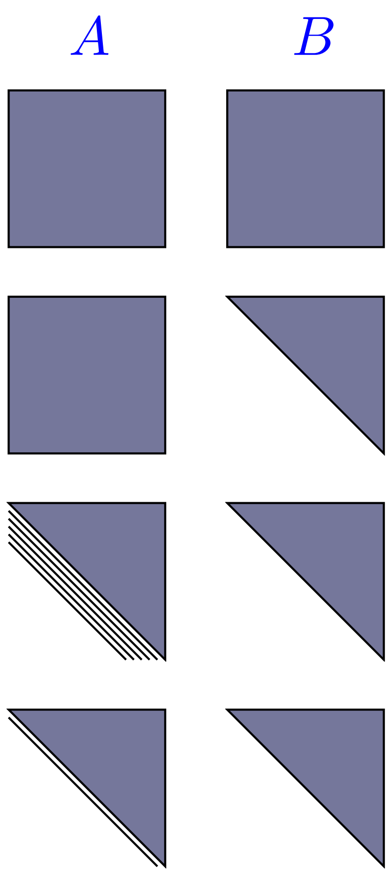
$$(A_i - \lambda B_i) \longrightarrow$$



Three ingredients of implicit QZ:

- initial reduction to Hessenberg-triangular form;
- deflation;
- QZ iterations $=$ bulge chasing.

Hessenberg-Triangular Reduction



original matrix pencil $A - \lambda B$

blocked QR factorization of B and update of A

blocked reduction to block Hessenberg-triangular form

reduction to Hessenberg-triangular form
based on pipelined Givens rotations

(Dackland/Kågström '99)

Up to three times faster than LAPACK's `DGGHRD`.

Deflation I: Small Subdiagonal Entry in A

$$\begin{bmatrix} a & a & a & a & a & a \\ a & a & a & a & a & a \\ & a & a & a & a & a \\ & & \varepsilon & a & a & a \\ & & & a & a & a \\ & & & & a & a \end{bmatrix} - \lambda \begin{bmatrix} b & b & b & b & b & b \\ & b & b & b & b & b \\ & & b & b & b & b \\ & & & b & b & b \\ & & & & b & b \\ & & & & & b \end{bmatrix}$$

Deflation I: Small Subdiagonal Entry in A

$$\begin{bmatrix} a & a & a & a & a & a \\ a & a & a & a & a & a \\ & a & a & a & a & a \\ & & \varepsilon & a & a & a \\ & & & a & a & a \\ & & & & a & a \end{bmatrix} - \lambda \begin{bmatrix} b & b & b & b & b & b \\ & b & b & b & b & b \\ & & b & b & b & b \\ & & & b & b & b \\ & & & & b & b \\ & & & & & b \end{bmatrix}$$

considerably small if $|a_{j+1,j}| \leq \mathbf{u} \cdot (|a_{jj}| + a_{j+1,j+1}|))$

Deflation I: Small Subdiagonal Entry in A

$$\left[\begin{array}{ccc|ccc} a & a & a & a & a & a \\ a & a & a & a & a & a \\ & a & a & a & a & a \\ \hline & & 0 & a & a & a \\ & & & a & a & a \\ & & & & a & a \end{array}\right] - \lambda \left[\begin{array}{ccc|ccc} b & b & b & b & b & b \\ & b & b & b & b & b \\ & & b & b & b & b \\ \hline & & & b & b & b \\ & & & & b & b \\ & & & & & b \end{array}\right]$$

$\Rightarrow$ generalized eigenvalue problem deflated into two smaller ones.

Deflation II: Small Diagonal Entry in B

$$\begin{bmatrix} a & a & a & a & a & a \\ a & a & a & a & a & a \\ & a & a & a & a & a \\ & & a & a & a & a \\ & & & a & a & a \\ & & & & a & a \end{bmatrix} - \lambda \begin{bmatrix} b & b & b & b & b & b \\ & b & b & b & b & b \\ & & \varepsilon & b & b & b \\ & & & b & b & b \\ & & & & b & b \\ & & & & & b \end{bmatrix}$$

Deflation II: Small Diagonal Entry in B

$$\begin{bmatrix} a & a & a & a & a & a \\ a & a & a & a & a & a \\ & a & a & a & a & a \\ & & a & a & a & a \\ & & & a & a & a \\ & & & & a & a \end{bmatrix} - \lambda \begin{bmatrix} b & b & b & b & b & b \\ & b & b & b & b & b \\ & & \varepsilon & b & b & b \\ & & & b & b & b \\ & & & & b & b \\ & & & & & b \end{bmatrix}$$

considerably small if $|b_{jj}| \leq \mathbf{u} \cdot (|b_{j-1,j}| + b_{j,j+1}|))$

Deflation II: Small Diagonal Entry in B

$$\begin{bmatrix} a & a & a & a & a & a \\ a & a & a & a & a & a \\ & a & a & a & a & a \\ & & a & a & a & a \\ & & & a & a & a \\ & & & & a & a \end{bmatrix} - \lambda \begin{bmatrix} b & b & b & b & b & b \\ & b & b & b & b & b \\ & & 0 & b & b & b \\ & & & b & b & b \\ & & & & b & b \\ & & & & & b \end{bmatrix}$$

Deflation II: Small Diagonal Entry in B

$$
\begin{bmatrix}
a & a & a & a & a & a \\
0 & a & a & a & a & a \\
 & a & a & a & a & a \\
 & & a & a & a & a \\
 & & & a & a & a \\
 & & & & a & a
\end{bmatrix}
- \lambda
\begin{bmatrix}
0 & b & b & b & b & b \\
 & b & b & b & b & b \\
 & & b & b & b & b \\
 & & & b & b & b \\
 & & & & b & b \\
 & & & & & b
\end{bmatrix}
$$

.. by a sequence of Givens rotations ..

Deflation II: Small Diagonal Entry in B

$$\left[\begin{array}{c|ccccc} a & a & a & a & a & a \\ \hline 0 & a & a & a & a & a \\ & a & a & a & a & a \\ & & a & a & a & a \\ & & & a & a & a \\ & & & & a & a \end{array}\right] - \lambda \left[\begin{array}{c|ccccc} 0 & b & b & b & b & b \\ \hline & b & b & b & b & b \\ & & b & b & b & b \\ & & & b & b & b \\ & & & & b & b \\ & & & & & b \end{array}\right]$$

$\Rightarrow$ one eigenvalue $\lambda = \infty$ deflated.

Implicit QZ Iteration

Assumption: B is nonsingular.

Goal: Drive subdiagonal entries of A to ε.

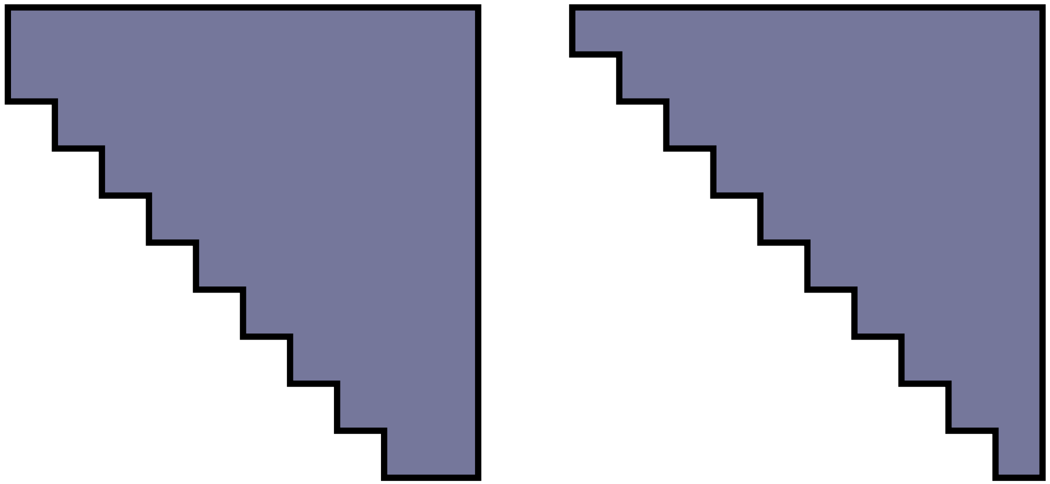
Shift polynomial: Define shifts σ_1, σ_2 as generalized eigenvalues of the bottom right 2×2 subpencil of $A - \lambda B$. Let

$$x = (AB^{-1} - \sigma_1 I_n)(AB^{-1} - \sigma_2 I_n)e_1,$$

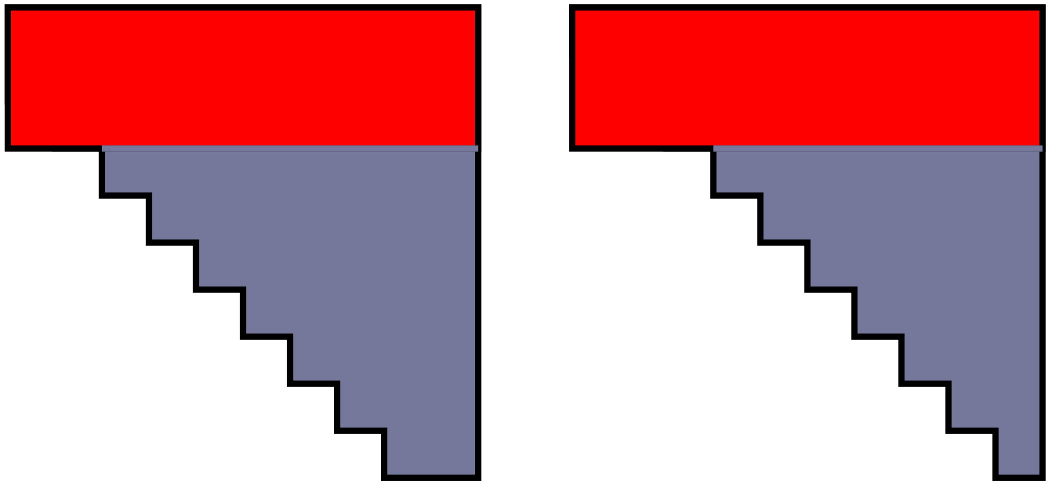
and Q such that $Q^T x = \alpha e_1$.

QZ iteration: Reduce $Q^T A - \lambda Q^T B$ back to Hessenberg-triangular form.

Bulge Chasing

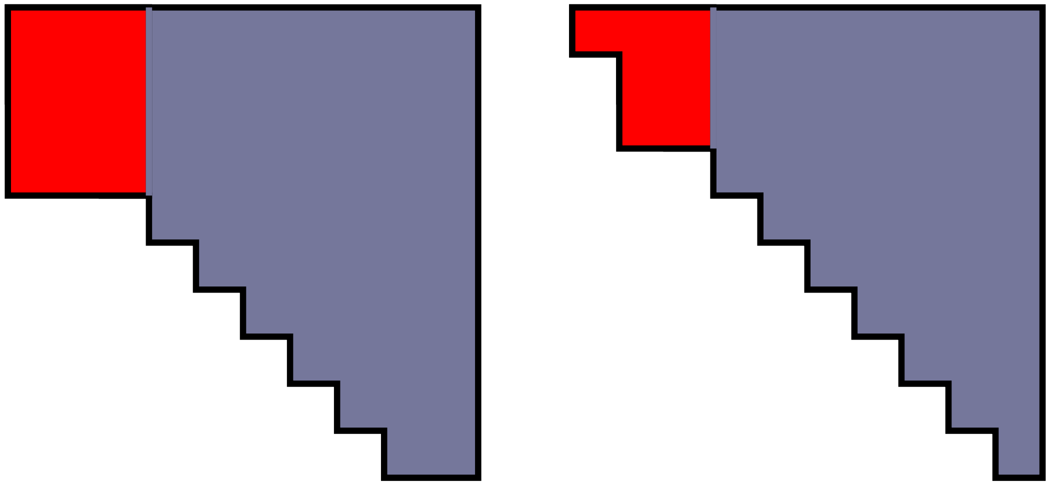


Bulge Chasing



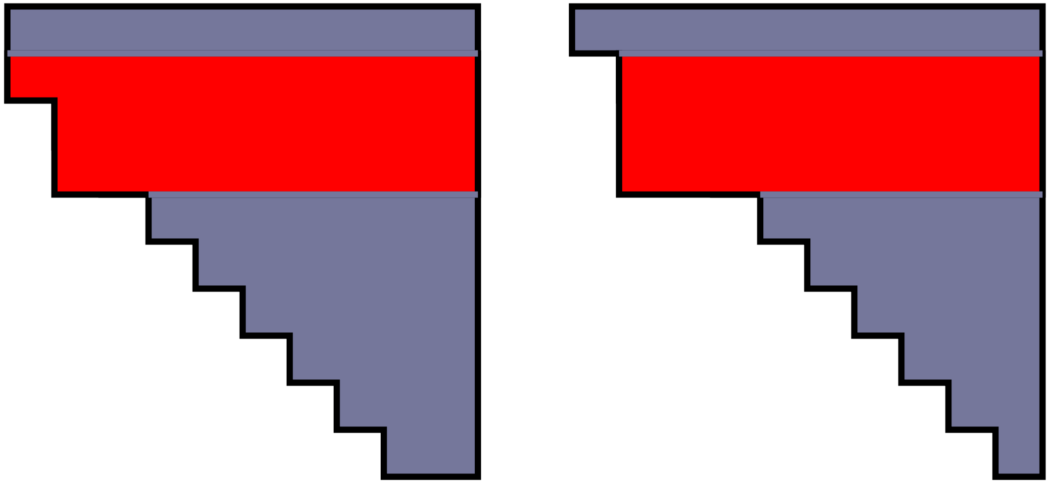
..apply Q^T from the left..

Bulge Chasing



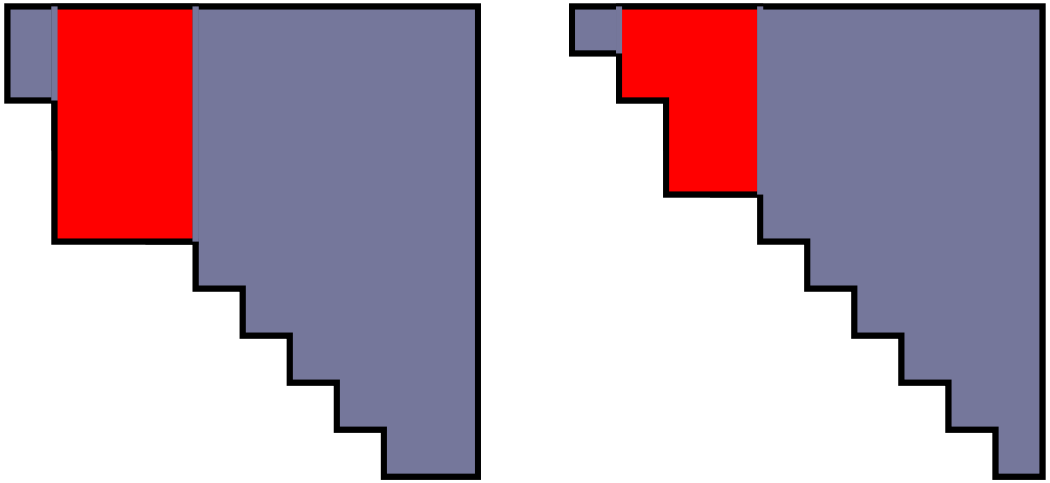
..reduce 1^{st} column of B by (opposite) Householder from the right..

Bulge Chasing



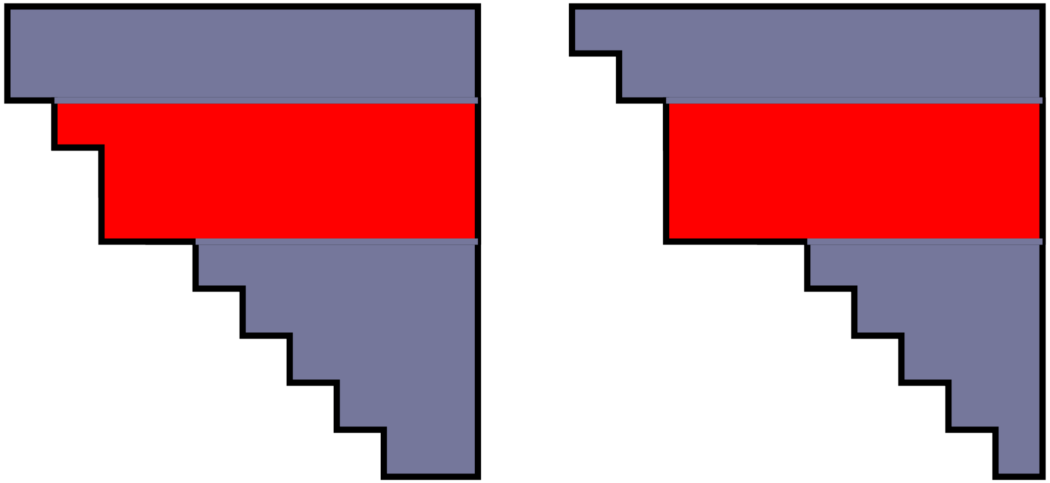
..reduce 1^{st} column of A by Householder from the left..

Bulge Chasing

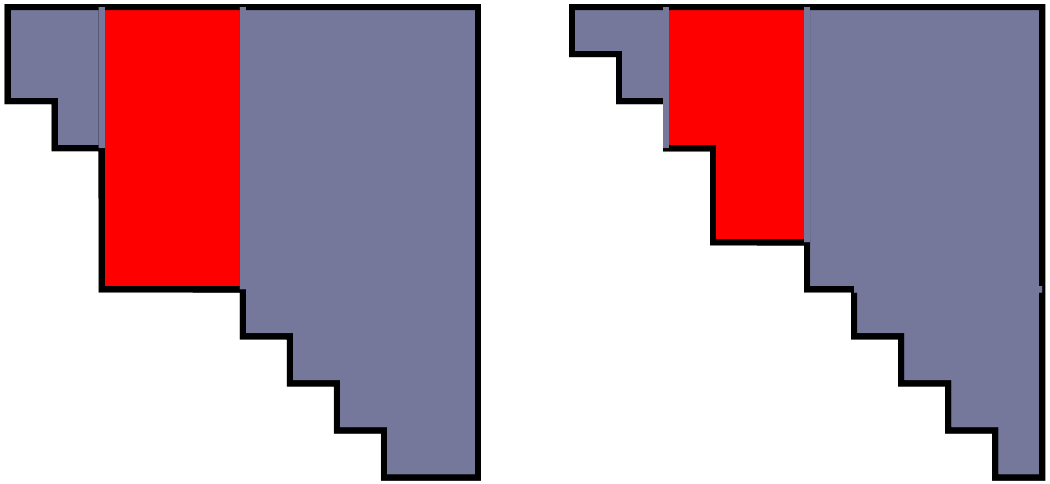


..reduce 2^{nd} column of B by (opposite) Householder from the right..

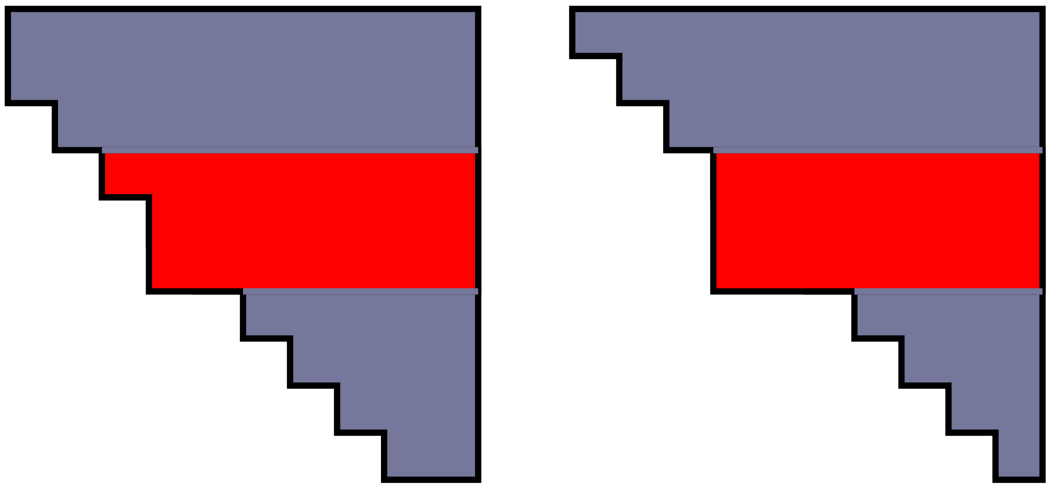
Bulge Chasing



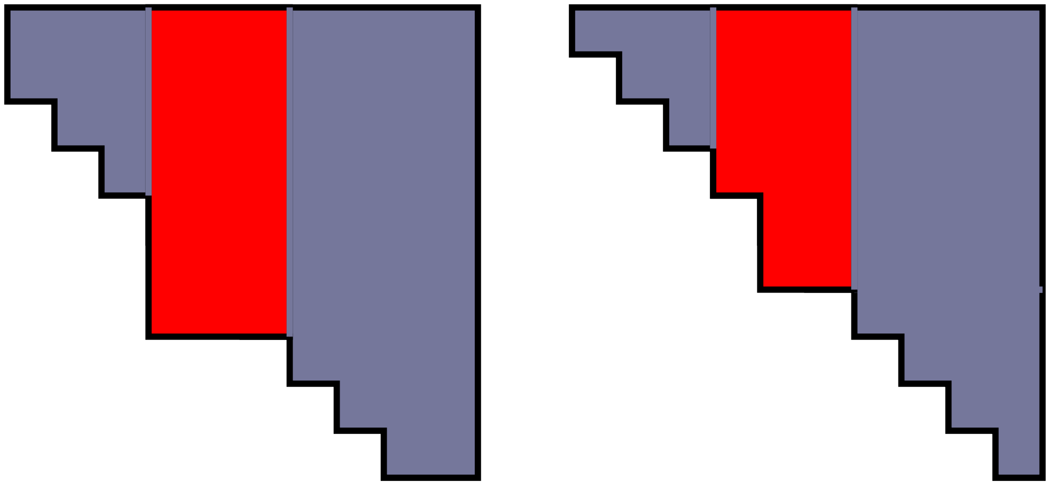
Bulge Chasing



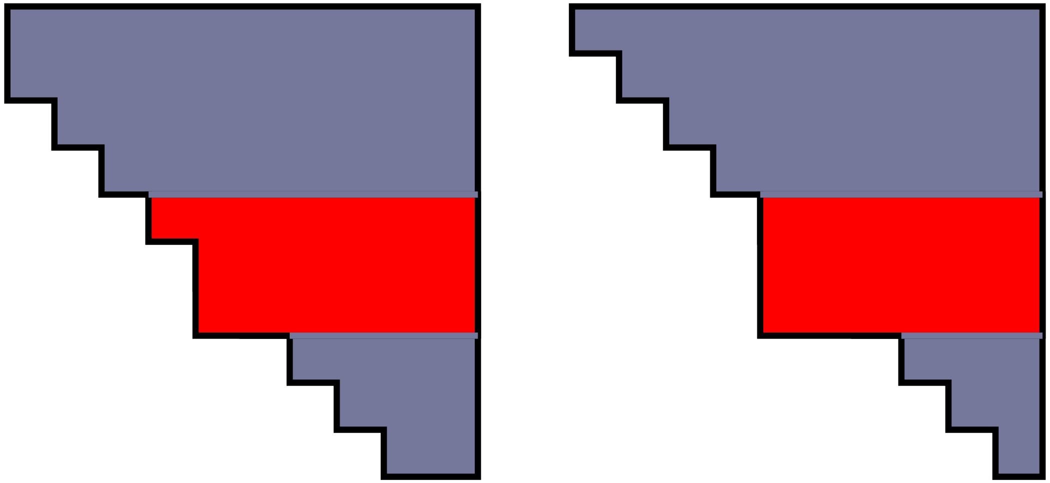
Bulge Chasing



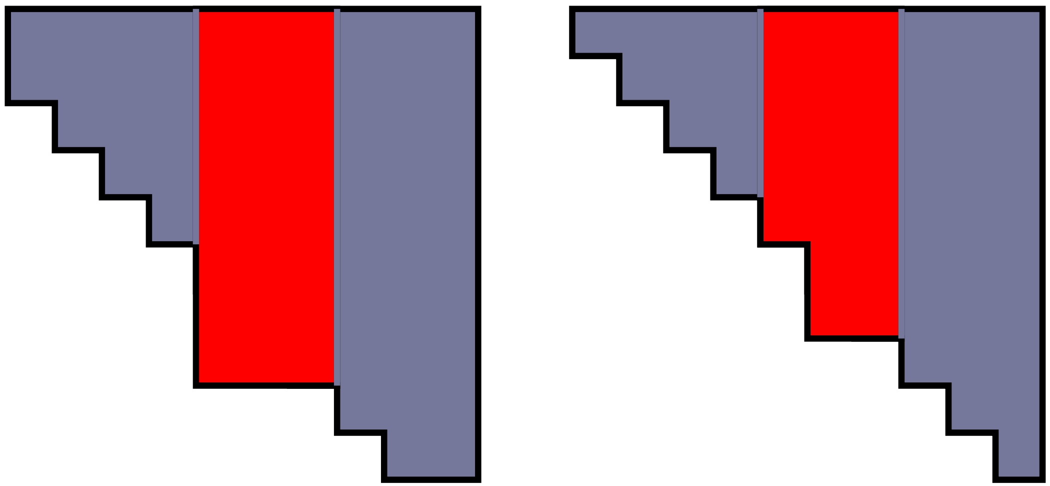
Bulge Chasing



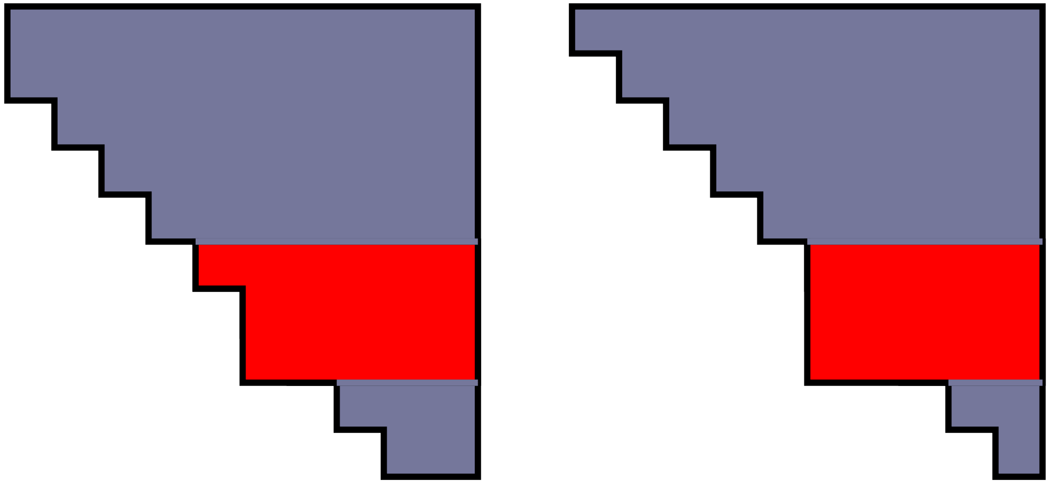
Bulge Chasing



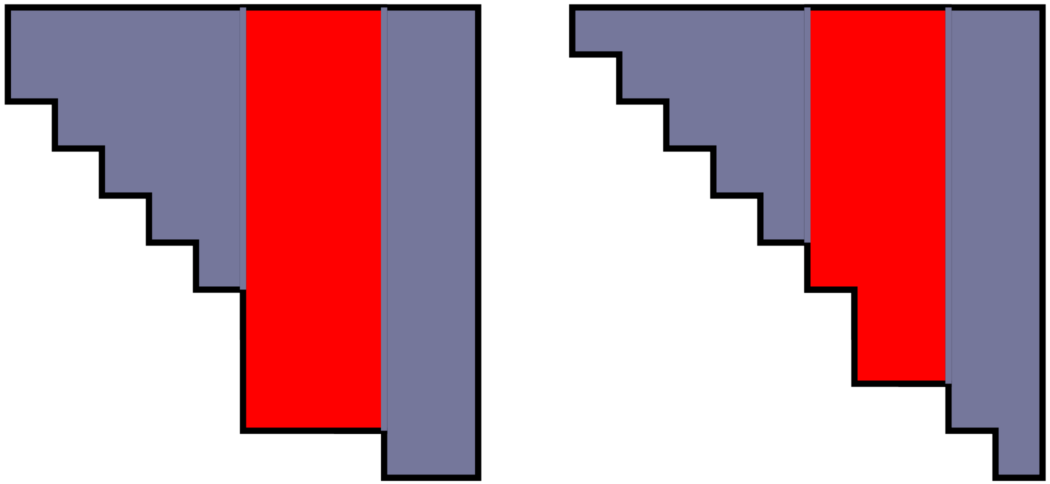
Bulge Chasing



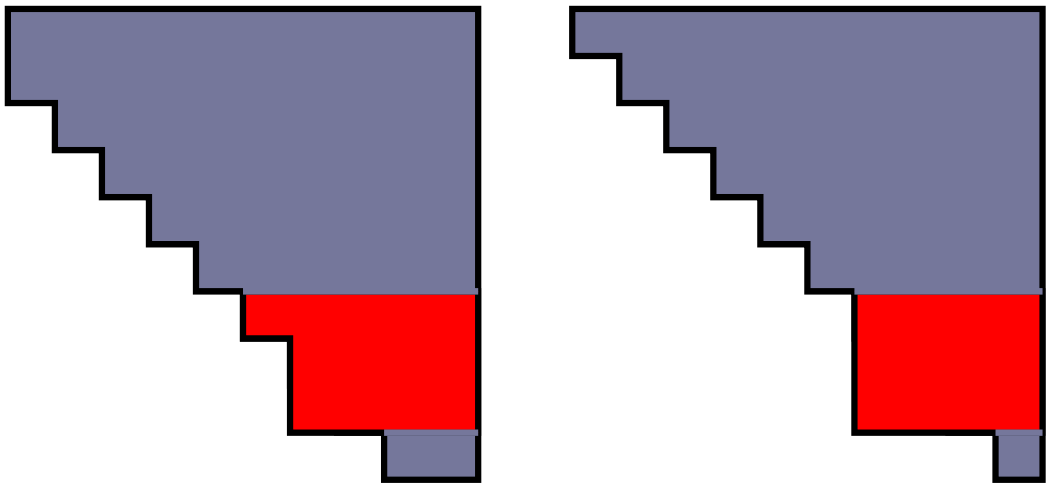
Bulge Chasing



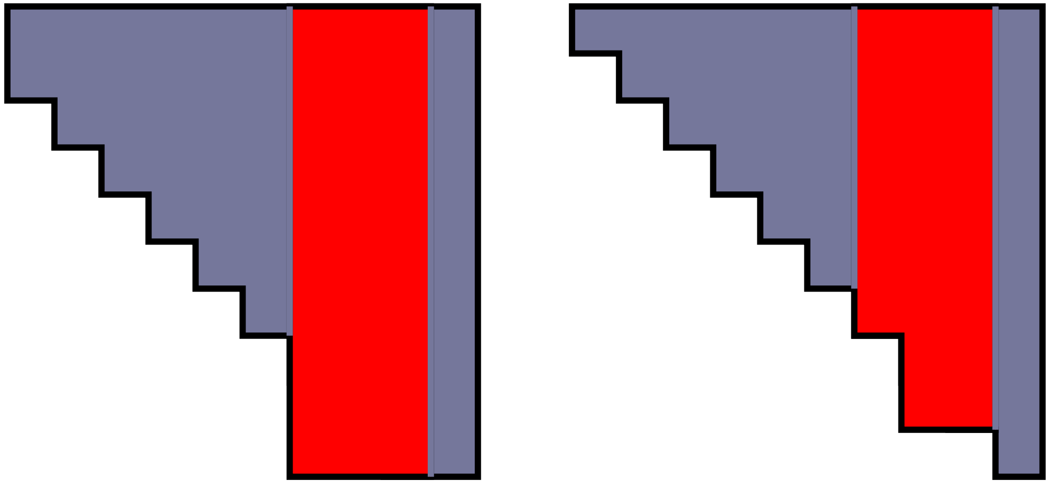
Bulge Chasing



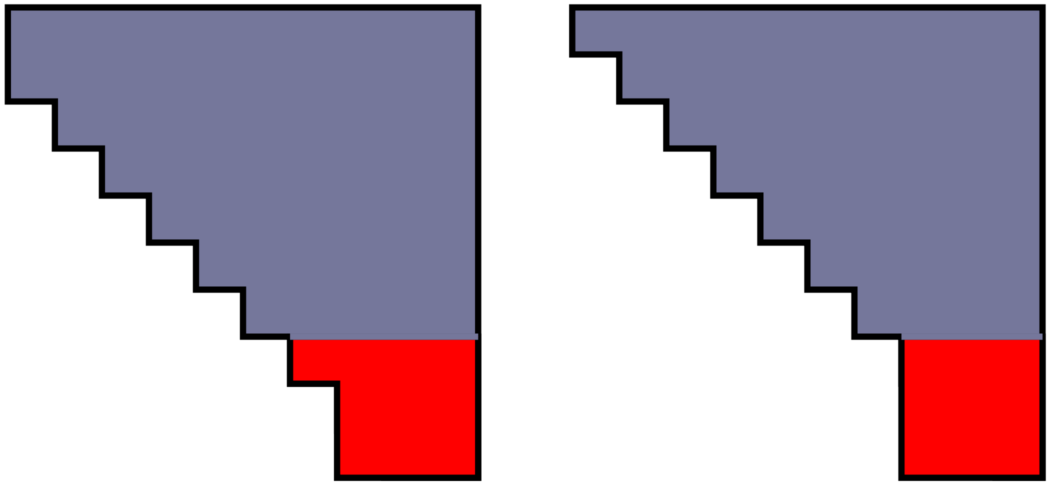
Bulge Chasing



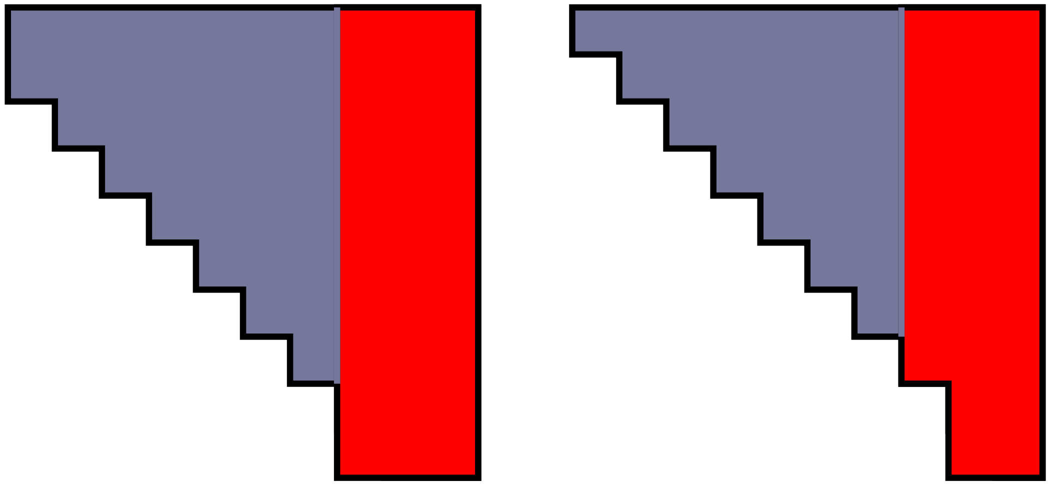
Bulge Chasing



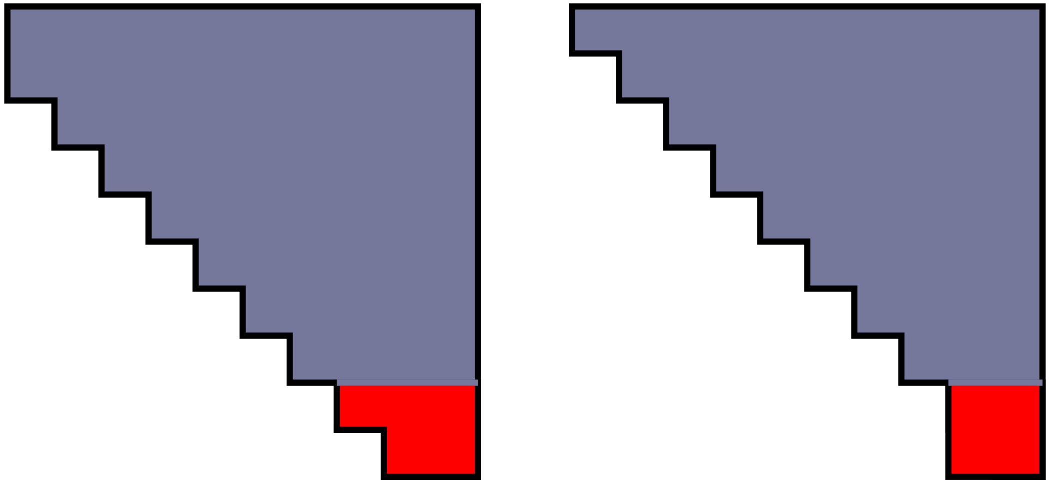
Bulge Chasing



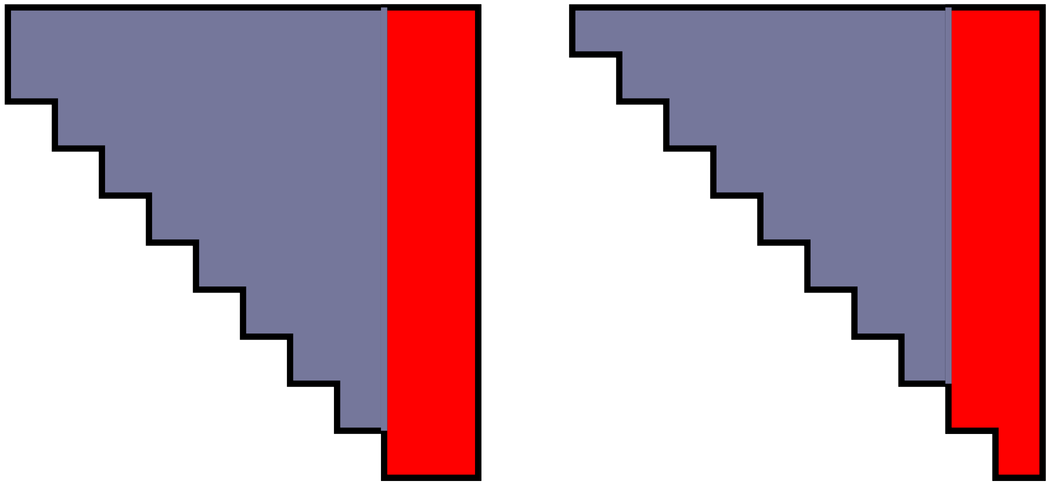
Bulge Chasing



Bulge Chasing



Bulge Chasing



Bulge Chasing

- One QZ iteration requires $\mathcal{O}(n^2)$ flops.
- On average, roughly two QZ iterations are necessary for deflating a gen. eigenvalue (typically at the bottom right corner).
- High memory access/computation ratio and poor memory access pattern $\Rightarrow$ poor performance!

Bulge Chasing

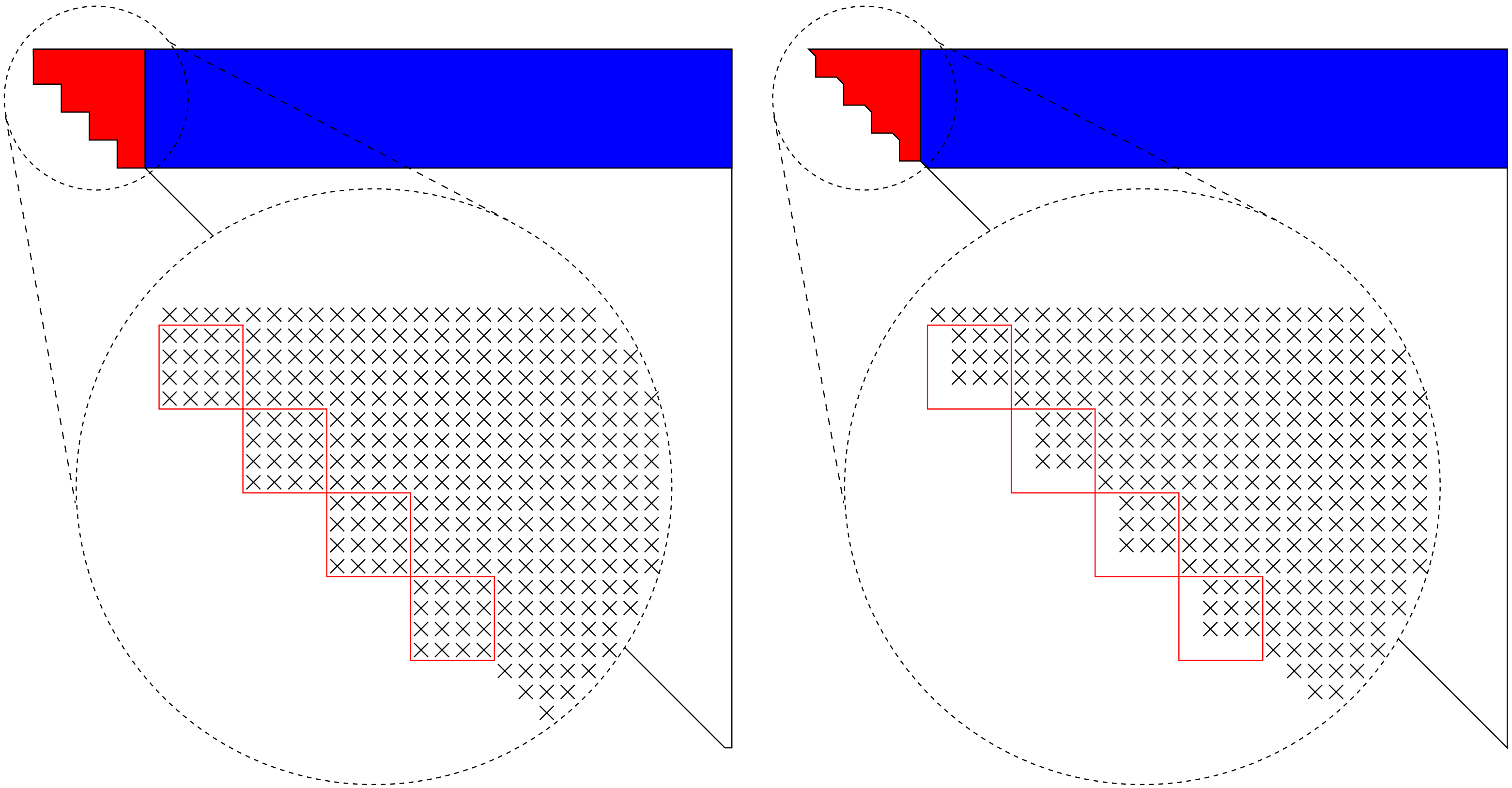
- One QZ iteration requires $\mathcal{O}(n^2)$ flops.
- On average, roughly two QZ iterations are necessary for deflating a gen. eigenvalue (typically at the bottom right corner).
- High memory access/computation ratio and poor memory access pattern $\Rightarrow$ poor performance!

Remedy: Use more shifts per iteration.

But: Large bulge sizes lead to shift blurring phenomena and loss of convergence(Watkins '96, Kressner '04).

Use tightly coupled chain of small bulges instead. Based on ideas of Braman/Byers/Mathias '02, Lang '97, and many others for the QR algorithm.

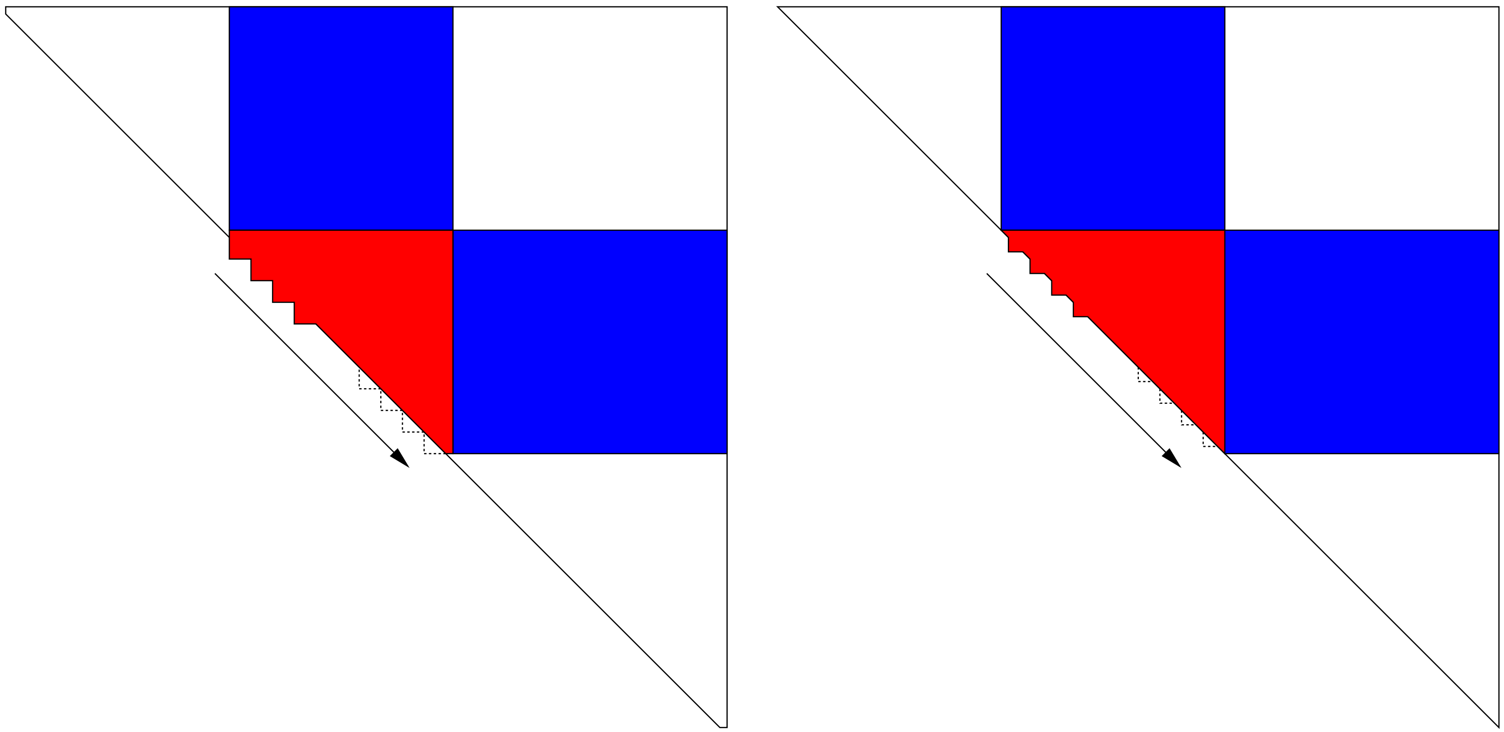
Multishift QZ: Introducing a Chain of Bulges



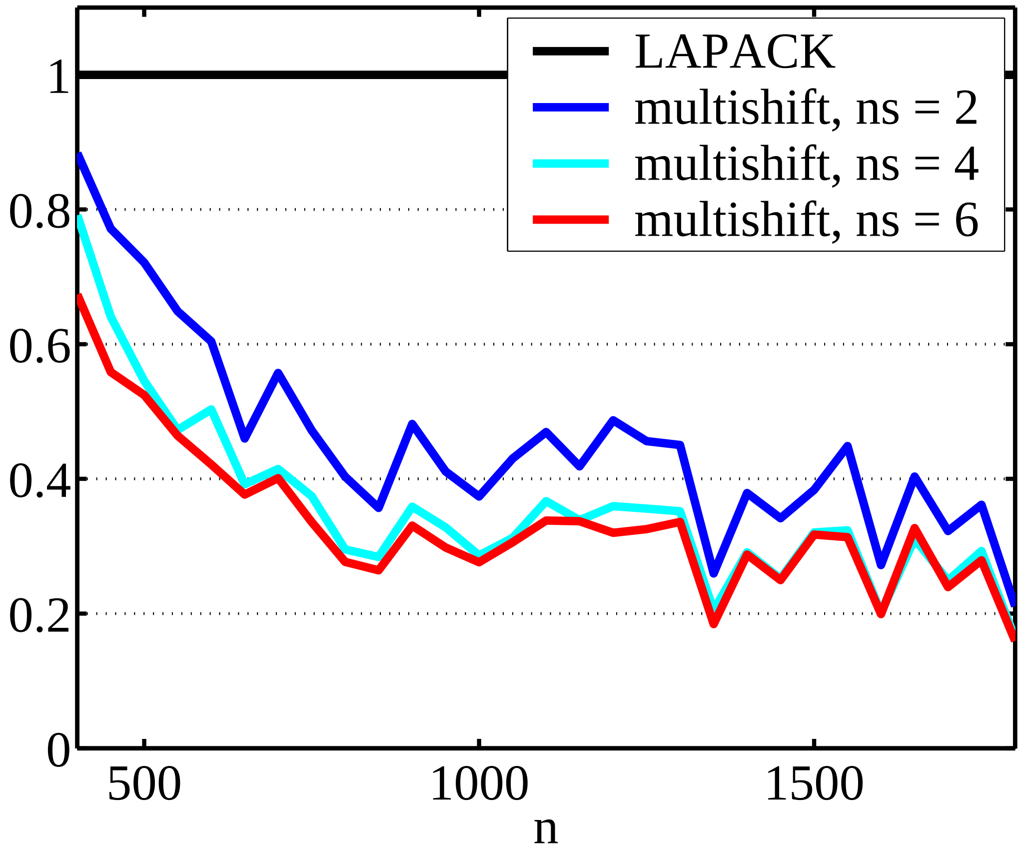
Red area: Updated during introduction.

Blue area: Updated after introduction via matrix-matrix-mult.

Multishift QZ: Chasing a Chain of Bulges



Performance of the QZ algorithm

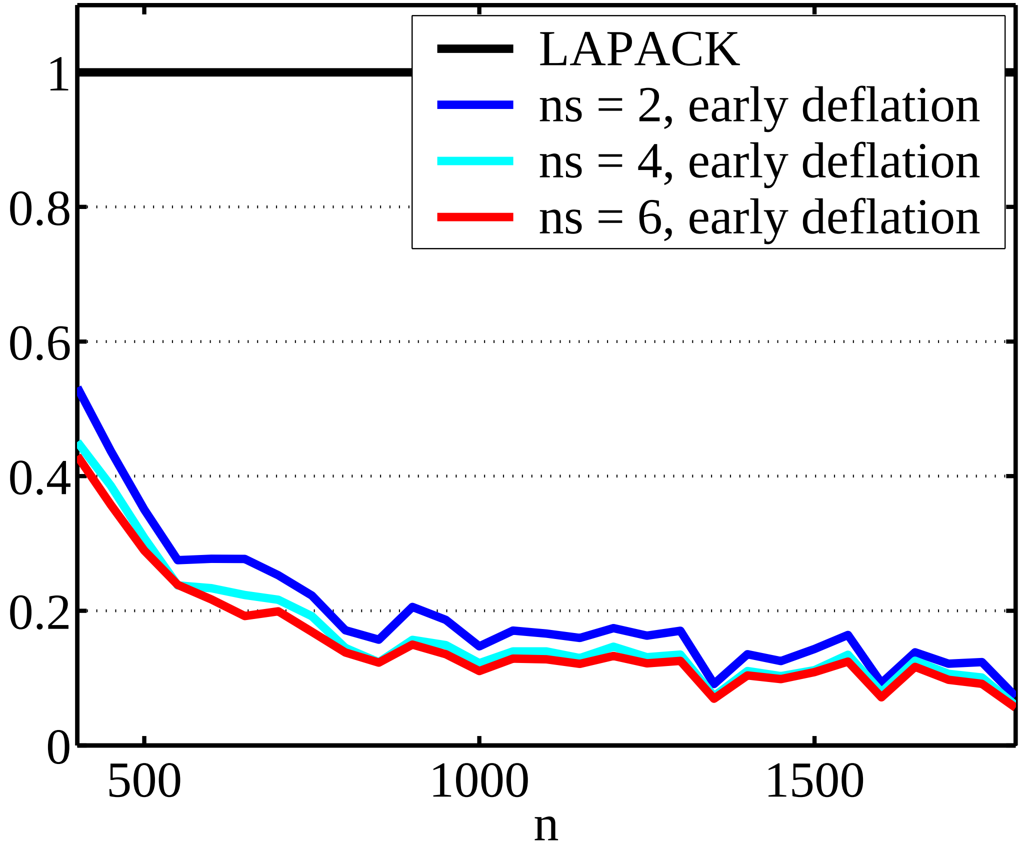


shifts/bulge: $n_s \in \{2, 4, 6\}$

shifts/QZ iteration: $m = 60$

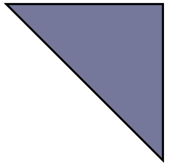
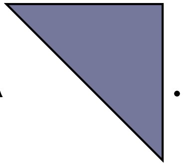
Aggressive Early Deflation

Based on work by Braman/Byers/Mathias ’02 for the QR algorithm, a more effective deflation strategy can be used to accelerate convergence of the QZ algorithm.



Computation of Deflating Subspaces

Output of QZ algorithm: Orthogonal matrices Q, Z such that

$$Q^T AZ - \lambda Q^T BZ =$$  $$- \lambda$$  .

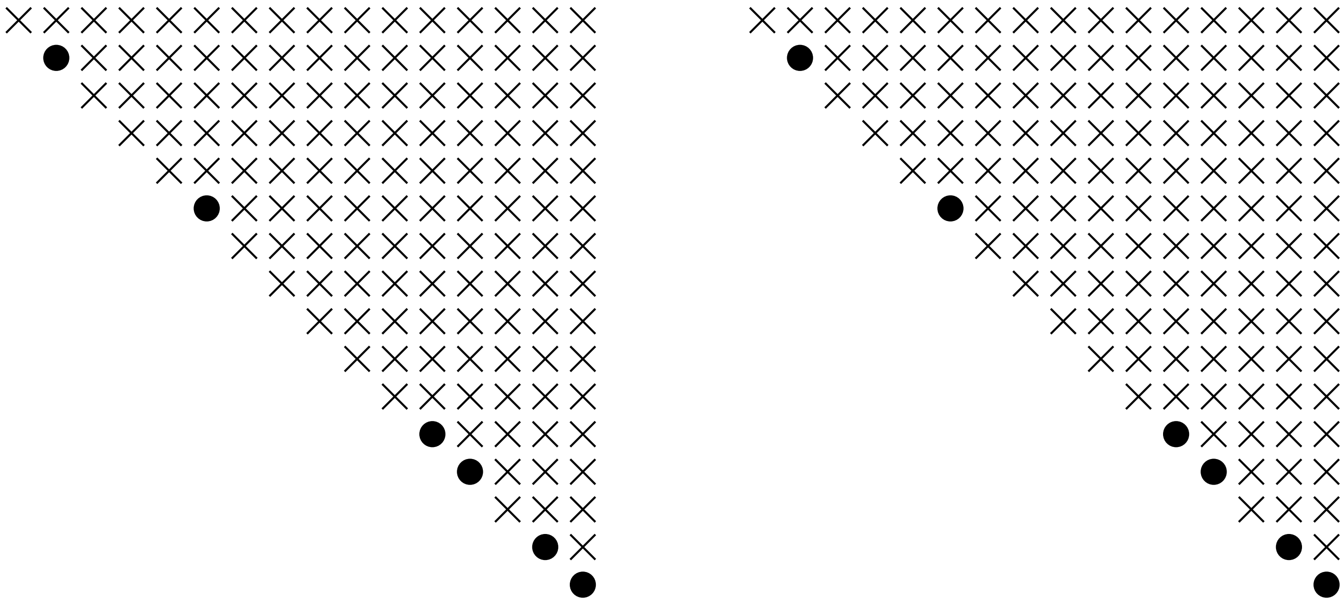
First k columns of Q and Z span pair of deflating subspaces belonging to gen. eigenvalues of $k \times k$ leading principal subpencil.

To compute other deflating subspaces, gen. eigenvalues must be reordered.

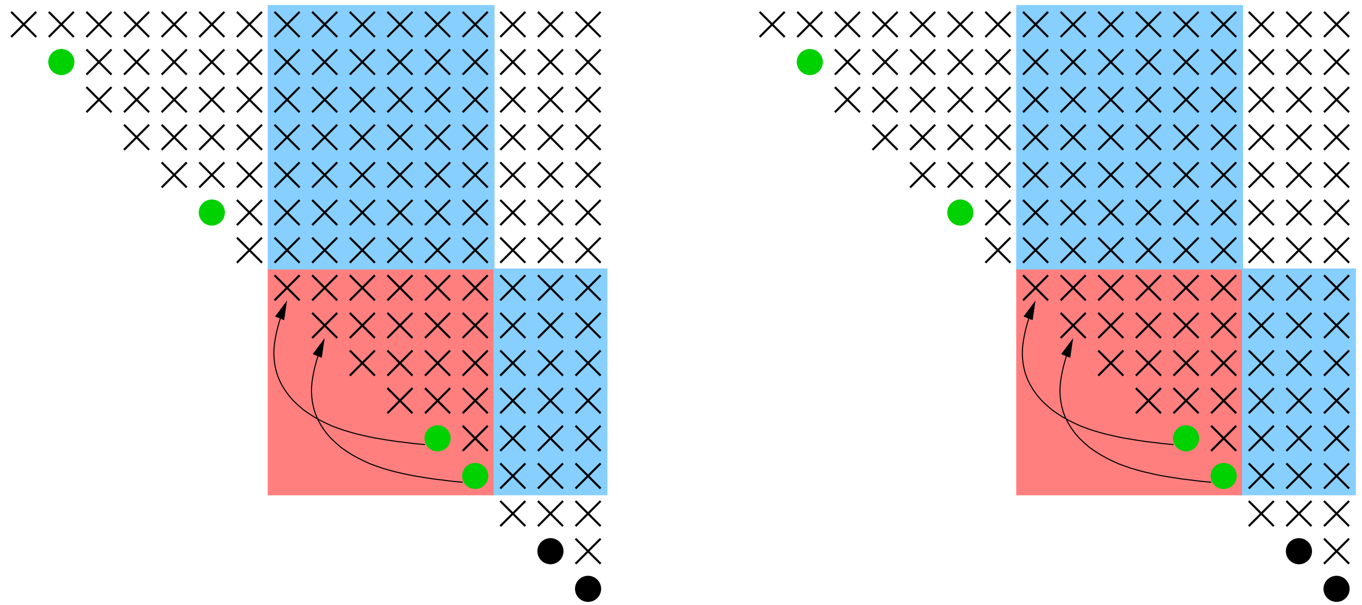
Van Dooren '82, Kågström '93, Kågström/Poromaa '96, propose to reorder gen. eigenvalues in a bubble sort-like fashion.

Again: High memory access/computation ratio and poor memory access pattern $\Rightarrow$ poor performance!

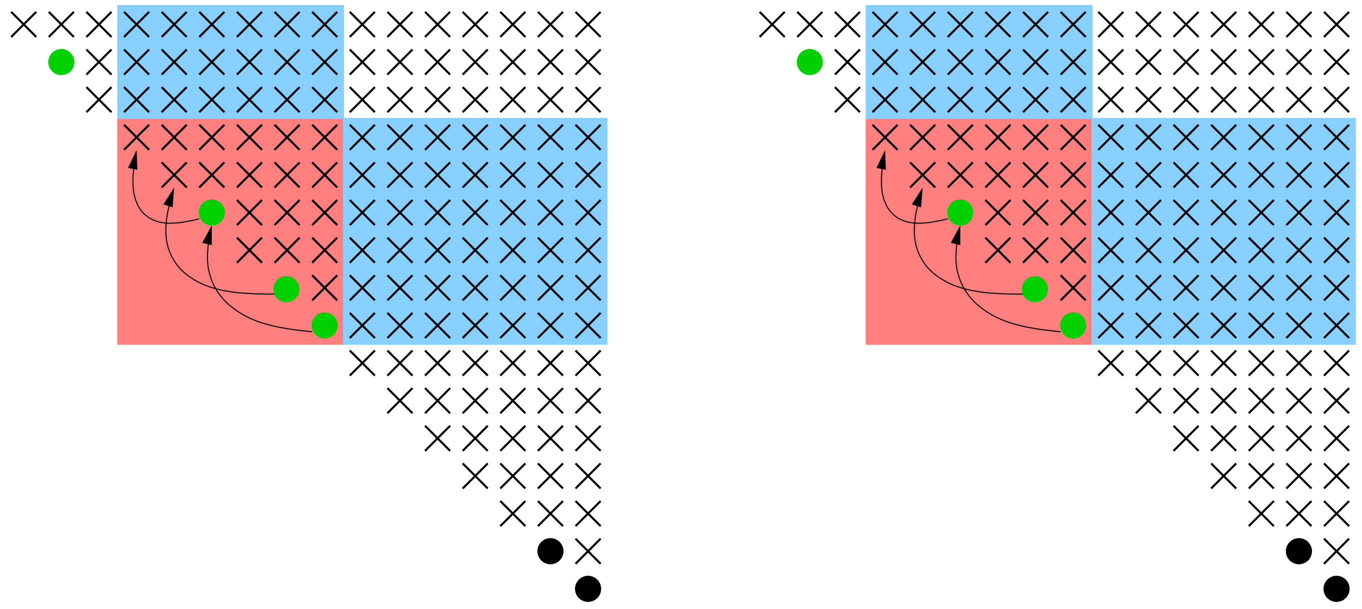
Block Algorithm for Reordering Eigenvalues



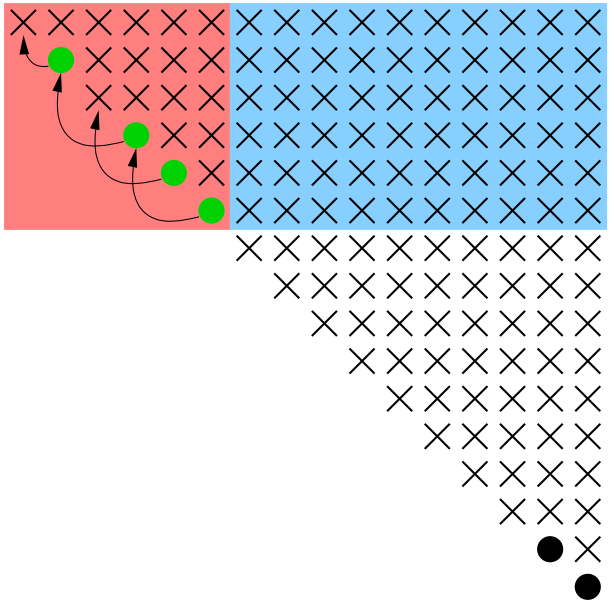
Block Algorithm for Reordering Eigenvalues

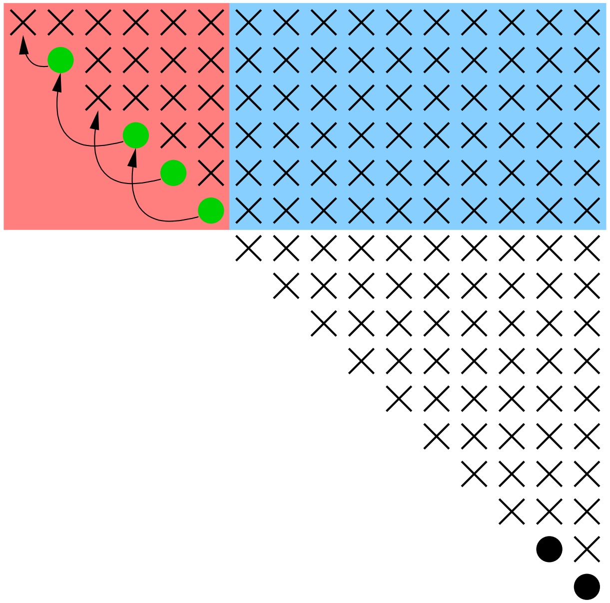


Block Algorithm for Reordering Eigenvalues



Block Algorithm for Reordering Eigenvalues





Performance of Block Reordering Algorithm

For standard eigenvalue problem:

n	sel.	LAPACK	new	ratio
500	5%	0.25	0.09	36%
500	25%	0.75	0.25	33%
500	50%	0.81	0.33	40%
1000	5%	2.87	0.60	21%
1000	25%	8.40	1.57	19%
1000	50%	10.08	2.10	21%
1500	5%	9.46	1.69	18%
1500	25%	30.53	4.88	16%
1500	50%	35.93	6.55	18%

Concluding Remarks

Work under progress:

- Integration of described algorithms into new release of LAPACK.
- ScaLAPACK-like parallel implementation of QZ algorithm (Björn Adlerborn, Univ. Umeå)

Concluding Remarks

Work under progress:

- Integration of described algorithms into new release of LAPACK.
- ScaLAPACK-like parallel implementation of QZ algorithm (Björn Adlerborn, Univ. Umeå)

For more details, see:

- Adlerborn/Dackland/Kågström: Parallel and blocked algorithms for reduction of a regular matrix pair to Hessenberg-triangular and generalized Schur forms. PARA2002, Springer-Verlag, LNCS, Vol. 2367, pp 319–328.
- Kressner/Kågström: Multishift variants of the QZ algorithm with aggressive early deflation. In preparation, 2004.
- Kressner: Numerical algorithms and software for general and structured eigenvalue problems. PhD thesis, 2004.